

Attorney Docket # 4925-87

Patent

AF/2700
2654
#14
a
4/28/04

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of

Riku SUOMELA et al.

Serial No.: 09/740,277

Filed: December 19, 2000

For: A Method For Activating Context Sensitive
Speech Recognition In A Terminal

Examiner: Nolan, Daniel A.
Group Art: 2654

I hereby certify that this correspondence is being
deposited with the United States Postal Service with
sufficient postage as first class mail in an envelope
addressed to: Commissioner for Patents, P.O. Box 1450,
Alexandria, VA 22313-1450, on

April 26, 2004
(Date of Deposit)

Alfred W. Froeblich

Name of applicant, assignee or Registered Representative

Alfred W. Froeblich
Signature

April 26, 2004
Date of Signature

Mail Stop **Appeal Brief - Patents**
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

RECEIVED

APR 30 2004

Technology Center 2600

APPEAL BRIEF

SIR:

This is an appeal, pursuant to 37 C.F.R. §1.192(a) from the decision of the Examiner in the above-identified application, as set forth in the Final Office Action wherein the Examiner finally rejected appellant's claims. The rejected claims are reproduced in the Appendix A attached hereto. A Notice of Appeal was filed on February 26, 2004. This Appeal Brief is being submitted in triplicate.

The fee of \$330.00 for filing an Appeal Brief (Large Entity) pursuant to 37 C.F.R. §1.17(f) is submitted herewith. Any additional fees or charges in connection with this application may be charged to our Patent and Trademark Office Deposit Account No. 03-2412.

04/29/2004 WAFSAW1 00000011 09740277

01 FC:1402

330.00 0P

REAL PARTY IN INTEREST

The assignee, Nokia Corporation, of applicants, Riku Suomela and Juha Lehtikoinen, is the real party of interest in the above-identified U.S. Patent Application.

RELATED APPEALS AND INTERFERENCES

There are no other appeals and/or interferences related to the above-identified application at the present time.

STATUS OF CLAIMS

The application was filed with claims 1-46. Claims 1, 2, 7, 11-14, 28, 34, 37-39, and 41-46 were amended and claims 47-49 were added by an amendment filed on September 5, 2003. Claims 1-49 have been finally rejected. Claims 1-49 are on appeal.

STATUS OF AMENDMENTS

An Amendment was filed on January 23, 2004 subsequent to the Final Office Action. In response, the Examiner issued an Advisory Action dated February 11, 2004 in which he refused to enter the Amendment filed on January 23, 2004. However, the Examiner did indicate in the Advisory Action that the objections to the specification and the objections to the claims were overcome by the Amendment. Accordingly, it is respectfully submitted that the amendment filed on January 23, 2004 now be entered, as materially reducing or simplifying the issues for appeal. The claims 1-49 listed in Appendix A show the changes made by the Amendment filed on January 23, 2004.

SUMMARY OF THE INVENTION

Appellants' invention relates to a method and apparatus for activating speech recognition in a terminal. As discussed in the Background section of the specification, speech recognition is a high consumer of a terminal's power and processing time (see page 2, lines 11-13 of the present specification). Accordingly, it is not practical to have speech recognition active at all times in a device such as a wireless terminal or mobile phone which has limited power and processing capacity (page 2, lines 15-17). It is also impractical to require manual activation of speech recognition in the device because the user must divert attention to the device to perform the additional step of activating the speech recognition, thus detracting from the advantages of speech recognition (page 2, line 18 to page 3, line 4).

The present invention overcomes the above-described problems by automatically activating speech recognition at a terminal in response to the detection of an event (page 4, lines 7-10). The terminal maintains the speech recognition in an activated state for a speech recognition time period. During this speech recognition time period, the terminal determines whether a voiced command or a command input by the primary input of the terminal is received (page 4, lines 12-13). After the speech recognition time period has elapsed, the speech recognition in the terminal is deactivated. If the second command is not received during the speech recognition time period, it may be received thereafter only through the primary input (page 4, lines 13-15).

ISSUES

1. Whether claims 1-17, 27-32, 40-41, and 47-49 are patentable under 35 U.S.C. §103 over U.S. Patent No. 6,012,030 (French-St. George) (hereafter referred to as 'French') in view of U.S. Patent No. 4,481,384 (Matthews)?

2. Whether claims 18-26, 33-39, and 42-46 are patentable under 35 U.S.C. §103 as unpatentable over French and Matthews and further in view of U.S. Patent No. 6,377,793 (Jenkins)?

GROUPING OF CLAIMS

The pending claims are 1-49, of which claims 1, 27, 40 are independent. The claims are grouped as follows:

Group I -- claims 1-17, 27-32, 40-41, and 47-49, which stand or fall together.

Group II -- claims 18-26, 33-39, and 42-46, which stand or fall together.

ARGUMENT

GROUP I (CLAIMS 1-17, 27-32, 40-41, AND 47-49)

Independent claim 1 is drawn to a method for activating speech recognition and recites automatically activating speech recognition in response to an event detected at the terminal. Independent claim 1 further recites that the speech recognition is deactivated if a second command has not been received and the speech recognition time period has elapsed and that the second command can be input only by the primary input after the speech recognition is deactivated.

It is respectfully submitted that independent claim 1 is allowable over French in view of Matthews because (1) there is no motivation to combine the deactivation after a period of time feature of Matthews with the method and device of French and (2) even if they were combined, French and Matthews fail to teach or suggest the claimed invention.

French discloses a device, system, and method for managing speech and audio prompts in response to a user's current interaction modality. According to French, a

communications device has a multimodal user interface, i.e., more than one mode of input and output (see col. 1, lines 61-65 and col. 4, lines 25-29). Furthermore, the state of the speech interface in the device of French is dynamically switchable from a foreground state to a background state, wherein speech interface, audio prompts, and speech-based error recovery are fully implemented in the foreground state, and speech prompts are replaced by a limited set of audio prompts (or earcons) and no speech-based error recovery is implemented in the background state (col. 3, lines 30-32, 40-42, and 45-47). French also teaches that the device may dynamically select the alternative states of speech interfaces, i.e., the foreground state or the background state, based on a user's input modality (col. 4, lines 33-35). For example, if the user's input modality is speech input, the foreground state is selected and if the user's input modality is non-speech input the background state is selected (col. 4, lines 9-14).

Accordingly, French discloses a device in which the speech interface is always activated in either the foreground state or the background state. As pointed out by the Examiner, the Background section of the French reference discloses the operation of speech recognition for a limited time window (col. 2, lines 47-50). However, French further explains in col. 2, line 55-col. 3 line 2, that the use of time windows is undesirable.

In any event, French fails to teach or suggest deactivating speech recognition if no commands have been received within a predetermined period after an event, as expressly recited in independent claim 1.

Matthews discloses a Voice recognizing telephone call denial system which allow a user to make a call if the user's password and voice are recognized. According to Matthews, once a subscriber gains access to a central office trunk, the central office portion of the trunk is terminated

in the present circuit and the PABX portion of the trunk is temporarily split and connected to a speech recognition circuitry (see col. 2, lines 36-41 of Matthews). The subscriber then speaks into the telephone handset, giving a password, his name, or another predetermined word (col. 2, lines 41-43). If the voice is not recognized, use of the trunk is denied or limited; if the voice is recognized, the system provides to the user a class of service associated with the user (col. 5, line 65 - col. 6, line 2). As indicated by the Examiner, col. 11, lines 66-68 of Matthews discloses that the speech recognition mode may be deactivated after a predetermined period of time. In other words, the system of Matthews does not wait indefinitely for a user to voice his or her password.

Regarding the first reason, a prior art reference must be considered in its entirety, i.e., as a whole, including portions that would lead away from the claimed invention. *W.L. Gore & Associates, Inc. v. Garlock, Inc.*, 721 F.2d 1540, 220 USPQ 303 (Fed. Cir. 1983), *cert. denied*, 469 U.S. 851 (1984). As discussed above, French discloses a multimodal user interface in which a user can either input a voiced command or input a command using an input. French discloses a system in which the user voices one of many available commands to a telecommunication system. Furthermore, the commands may be required at different times of operation. French discloses that the implementation of time windows for response in speech recognition is disadvantageous because users are not aware of how long they have to provide a voiced input (see col. 2, line 63 - col. 3, line 2). French overcomes these disadvantages of time windows by implementing a foreground state and a background state for speech recognition. To apply a time window for speech recognition to French would go against the teachings of French and therefore would not be implemented. Accordingly, the time window disclosed by Matthews would not be

implemented in the device and method of French because French teaches that such time windows are disadvantageous.

Even if the teaching of Matthews were combined with the teaching of French, the combined teachings still fail to teach or suggest the claimed invention. As described above, French discloses a device having a multimodal interface, wherein one of the interfaces is a speech recognition interface. According to French, a user can input one of a plurality of commands using a voiced command or through an input device. In contrast to French, Matthews is not inputting a command to be performed using voice recognition. Rather, Matthews uses voice recognition to record an utterance by a user to determine the identification of a user and the appropriate level of access of the user.

It is appropriate to use time windows in Matthews because the system should not wait indefinitely for an input by a user in the access mode. However, French teaches that speech recognition should always be available for receiving voiced commands to be performed. Since French teaches that speech recognition is always available for receiving voiced commands, the combined teaching of French and Matthews fails to teach or suggest the step of deactivating speech recognition at the terminal if it is determined that the second command is not received by the speech recognition or the primary input during a speech recognition time period, as expressly recited in independent claim 1.

In view of the above amendments and remarks, it is respectfully submitted that independent claim 1 is allowable over French in view of Matthews under 35 U.S.C. §103.

Independent claims 27 and 40 are directed to a terminal capable of speech recognition and a system for activating speech recognition in a terminal, respectively. Both of these

claims recite that speech recognition is activated in response to an event at the terminal and that the speech recognition is deactivated after a speech recognition time period has elapsed.

As stated above, French in view of Matthews fails to teach or suggest deactivating the speech recognition after a speech recognition time period because French teaches that the device is always available for receiving voiced commands. Accordingly, it is respectfully submitted that independent claims 27 and 40 are also allowable over French under 35 U.S.C. §103.

Dependent claims 2-26, 28-39, and 41-49, each being dependent on one of independent claims 1, 27, and 40, are deemed allowable for the same reasons expressed above with respect to independent claims 1, 27, and 40.

For the foregoing reasons, it is respectfully submitted that the combined teachings of fail to establish a *prima facie* case of obviousness with regard to the subject matter recited in claims 1-17, 27-32, 40-41, and 47-49. The Final Rejection of the claims in Group I should be reversed.

GROUP II (CLAIMS 18-26, 33-39, AND 42-46)

Furthermore, it is respectfully submitted that independent claims 18-26, 33-39, and 42-46 are allowable over French in view of Matthews and further in view of Jenkins. Jenkins relates to a device and method for recording messages at coordinate waypoints in a subscriber service. Although the subscriber device may be a wearable computer (see col. 5, lines 29-40), there is no teaching or suggestion for a context aware application. There is no teaching or suggestion in Jenkins that speech input recognition is activated by receipt of a context related event at the terminal. In col. 5, lines 31-41, Jenkins mentions devices which communicate with commercial wireless networks. Based on the context in which it is used, Jenkins is referring to

wireless communication networks. Accordingly, Jenkins fails to disclose local access point and a local area network, as recited in dependent claim 24.

Furthermore, Jenkins fails to teach or suggest "a wearable computer with a context aware application", as recited in dependent claims 19, 34, and 42, "an object in the environment has been selected" as recited in dependent claims 20, 36, and 42, "the notification is one of a phone call and a short message service message", as recited in dependent claims 23, 36, and 43, "connecting to one of a local access point and a local area network via short range radio technology", as recited in dependent claims 24, 37, and 44, "receiving information at the terminal from the computer environment of the terminal", as recited in dependent claims 25, 38, and 45, and "connecting to a site on the internet", as recited in dependent claims 26, 39, and 46. Accordingly, it is respectfully submitted that independent claims 18-26, 33-39, and 42-46 are each allowable for these additional reasons.

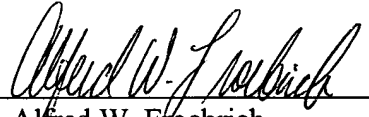
Dependent claims 47-49 each recite that the event which triggers voice recognition is either receipt of information at the terminal from the environment of the terminal or notification of an external event. The Examiner states that col. 6, lines 27-30 of French discloses this. However, the section of French referred to by the Examiner discloses that speech recognition is kept in the foreground state or switched to the foreground state in response to user input. It is respectfully submitted that user input is not information received at the terminal from the environment or notification of an external event, as recited in dependent claims 47-49. Accordingly, it is respectfully submitted that dependent claims 47-49 are allowable for these additional reasons.

For the foregoing reasons, it is respectfully submitted that the combined teachings of fail to establish a *prima facie* case of obviousness with regard to the subject matter recited in claims 18-26, 33-39, and 42-46. The Final Rejection of the claims in Group II should be reversed.

CONCLUSION

For the foregoing reasons, it is respectfully submitted that appellants' claims are not rendered obvious by and are, therefore, patentable over the art of record, and the Examiner's rejections should be reversed.

Respectfully submitted,
COHEN, PONTANI, LIEBERMAN & PAVANE

By 
Alfred W. Froeblich
Reg. No. 38,887
551 Fifth Avenue, Suite 1210
New York, New York 10176
Tel (212) 687-2770

Dated: April 26, 2004



APPENDIX

1. (previously presented) A method for activating speech recognition in a terminal, comprising the steps of:

- (a) detecting an event at the terminal;
- (b) performing a first command in response to the event of step (a);
- (c) automatically activating speech recognition at the terminal in response to said step (a);
- (d) determining whether a second command is received by one of the speech recognition and a primary input of the terminal during a speech recognition time period commenced upon a completion of said step (b);
- (e) deactivating the speech recognition at the terminal if it is determined that the second command is not received by the one of the speech recognition and a primary input of the terminal in said step (d) during the speech recognition time period;
- (f) determining whether the second command is received by the primary input after step (e); and
- (g) performing the second command received in one of said steps (d) and (f).

2. (previously presented) The method of claim 1, wherein said step (a) comprises detecting a use of a primary input of the terminal.

3. (original) The method of claim 1, wherein said step (c) further comprises determining a context in which speech recognition is activated and determining a word set of applicable commands in that context.

4. (original) The method of claim 3, wherein the word set determined in said step (c) comprises a default word set comprising commands that are applicable in all contexts.

5. (original) The method of claim 3, wherein said step (c) further comprises displaying at least a portion of the applicable commands of the word set.

6. (original) The method of claim 3, wherein said step (c) further comprises audibly outputting the applicable commands of the word set.

7. (previously presented) The method of claim 1, wherein said step (g) further comprises verifying that the second command received via speech recognition is correct.

8. (original) The method of claim 1, wherein said step (c) further comprises displaying at least a portion of the applicable commands of the word set.

9. (original) The method of claim 1, wherein said step (c) further comprises audibly outputting the applicable commands of the word set.

10. (original) The method of claim 1, wherein said step (d) further comprises receiving at least one second command via speech recognition during the speech recognition time period and saving said at least one second command in a command buffer.

11. (previously presented) The method of claim 10, wherein said step (g) comprises performing each command of said at least one second command in said command buffer.

12. (previously presented) The method of claim 11, further comprising the step of repeating said steps (c)-(g) in response to the command last performed in said step (g).

13. (previously presented) The method of claim 1, further comprising the step of repeating said steps (c)-(g) for the command last performed in said step (g).

14. (previously presented) The method of claim 11, further comprising the step of repeating said steps (c)-(g) in response to the last command performed by said step (g) if it is determined that the last command performed in said step (g) is an input defined to activate speech recognition.

15. (original) The method of claim 1, further comprising the step of determining whether the first command input in said step (a) is a command defined to activate speech recognition and wherein said steps (b) - (d) are performed only if it is determined that the first command performed in said step (a) is an action defined to activate speech recognition.

16. (original) The method of claim 1, wherein said step (a) comprises pressing a button.

17. (original) The method of claim 1, wherein said step (a) comprises pressing a button on a mobile phone.

18. (original) The method of claim 1, wherein said step (a) comprises pressing a button on a personal digital assistant.

19. (original) The method of claim 1, wherein the terminal is a wearable computer with a context-aware application and said step (a) comprises receiving information from the environment of the wearable computer.

20. (original) The method of claim 19, wherein the information is that an object in the environment has been selected.

21. (original) The method of claim 20, wherein the second command is an open command for accessing information about the selected object.

22. (original) The method of claim 1, wherein step (a) comprises receiving a notification from an external source.

23. (currently amended) The method of claim 22, wherein the notification is one of a phone call and a short message service message.

24. (original) The method of claim 1, wherein said step (a) comprises connecting to one of a local access point and a local area network via short range radio technology.

25. (original) The method of claim 1, wherein said step (a) comprises receiving information at the terminal from the computer environment of the terminal.

26. (original) The method of claim 25, wherein said step (a) comprises connecting to a site on the internet.

27. (currently amended) A terminal capable of speech recognition, comprising:

a central processing unit;

a memory unit connected to said central processing unit;

a primary input connected to said central processing unit for receiving inputted commands;

a secondary input connected to said central processing unit for receiving audible commands;

a speech recognition algorithm connected to said central processing unit for executing speech recognition; and

a primary control circuit connected to said central processing unit for
processing said inputted and audible commands, ~~and~~
activating speech recognition in response to an event for a speech
recognition time period, and
deactivating speech recognition after the speech recognition time period
has elapsed.

28. (previously presented) The terminal of claim 27, wherein said event
comprises a use of a primary input of the terminal.

29. (original) The terminal of claim 27, further comprising a word set database
connected to said central processing unit and a secondary control circuit connected to said central
processing unit for determining a context in which the speech recognition is activated and
determining a word set of applicable commands in said context from said word set database.

30. (original) The terminal of claim 29, further comprising a display for
displaying at least a portion of said word set.

31. (original) The terminal of claim 27, wherein said primary input comprises
buttons.

32. (original) The terminal of claim 31, wherein said terminal comprises a
mobile phone.

33. (original) The terminal of claim 31, wherein said terminal comprises a personal digital assistant.

34. (previously presented) The terminal of claim 27, wherein said terminal comprises a wearable computer with a context-aware application.

35. (original) The terminal of claim 34, wherein said means for activating speech recognition comprises means for activating speech recognition in response to a selection of an object in an environment of said wearable computer.

36. (currently amended) The terminal of claim 27, wherein said means for activating speech recognition comprises means for activating speech recognition in response to receiving notification of one of a phone call and a short message service message at said terminal.

37. (previously presented) The terminal of claim 27, wherein said means for activating speech recognition comprises means for activating speech recognition in response to connecting said terminal to one of a local access point and a local area network via short range radio technology.

38. (previously presented) The terminal of claim 27, wherein said means for activating speech recognition comprises means for activating speech recognition in response to receiving information at said terminal from a computer environment of said terminal.

39. (previously presented) The terminal of claim 38, wherein said means for activating speech recognition comprises means for activating speech recognition in response to connecting said terminal to a site on the internet.

40. (original) A system for activating speech recognition in a terminal, comprising:

a central processing unit;

a memory unit connected to said processing unit;

a primary input connected to said central processing unit for receiving inputted commands;

a secondary input connected to said central processing unit for receiving audible commands;

a speech recognition algorithm connected to said central processing unit for executing speech recognition; and

software means operative on the processor for

maintaining in said memory unit a database identifying at least one context related word set,

scanning for an event at the terminal,

performing a first command in response to the event,

activating speech recognition by executing said speech recognition algorithm for a speech recognition time period in response to detecting said event at said terminal,

deactivating speech recognition after the speech recognition time period has elapsed, and

performing a second command received during said speech recognition time.

41. (previously presented) The system of claim 40, wherein said event comprises a use of a primary input of the terminal.

42. (previously presented) The system of claim 40, wherein said means for activating speech recognition comprises means for activating speech recognition in response to a selection of an object in an environment of a wearable computer with a context-aware application.

43. (currently amended) The system of claim 40, wherein said means for activating speech recognition comprises means for activating speech recognition in response to receiving notification of one of a phone call and a short message service message at said terminal.

44. (previously presented) The system of claim 40, wherein said means for activating speech recognition comprises means for activating speech recognition in response to

connecting said terminal to one of a local access point and a local area network via short range radio technology.

45. (previously presented) The system of claim 40, wherein said means for activating speech recognition comprises means for activating speech recognition in response to receiving information at said terminal from a computer environment of said terminal.

46. (previously presented) The system of claim 45, wherein said means for activating speech recognition comprises means for activating speech recognition in response to connecting said terminal to a site on the internet.

47. (previously presented) The method of claim 1, wherein said step (a) comprises detecting one of receipt of information at the terminal from the environment of the terminal and notification of an external event.

48. (previously presented) The terminal of claim 27, wherein the event comprises one of receipt of information from the environment of the terminal and notification of an external event.

49. (previously presented) The system of claim 40, wherein the event comprises one of receipt of information from the environment of the terminal and notification of an external event.